IMPACT OF DIFFERENT MICROPHYSICS SCHEMES AND ADDITIONAL SURFACE OBSERVATIONS ON NEWS-E FORECASTS

Francesca M. Lappin[1], Dustan M. Wheatley[2,3], Kent H. Knopfmeier[2,3], and Patrick S. Skinner[2,3]

[1]National Weather Center Research Experiences for Undergraduates Program
Norman, Oklahoma
[2]Cooperative Institute for Mesoscale Meteorological Studies, The University of Oklahoma
Norman, Oklahoma
[3]NOAA/OAR/National Severe Storms Laboratory
Norman, Oklahoma

ABSTRACT

The National Weather Service issues hundreds of hazardous weather advisories, such as tornado warnings each year. However, the current warning paradigm has led to false alarm rates of 70% - 80% (Simmons and Sutter, 2011) and a stagnancy in warning lead times. In an attempt to mitigate these issues, the Warn-on-Forecast (WoF) project is developing a regional, high-resolution, storm-scale prediction system capable of predicting hazardous weather phenomena on the 0-3 h time scale. A prototype system, the NSSL Experimental Warn-on-Forecast System for ensembles (NEWS-e) combines, surface observations, radar data, and satellite retrievals in an ensemble data assimilation framework to create rapidly-updating probabilistic forecasts. Testing of the real-time configuration of the NEWS-e was completed during the 2017 Hazardous Weather Testbed Spring Forecast Experiment (HWT-SFE).

Two separate experiments were performed on three events to evaluate how changes in the configuration affect the performance of the NEWS-e. The first was a data denial experiment where the METAR surface observations were removed from the data assimilation. The second was a switch from the NSSL two-moment to the Thompson microphysics scheme. Quantitative metrics, namely the probability of detection (POD) and false alarm rate (FAR) for rotational objects, were computed to examine the overall performance of each NEWS-e forecast. The real-time run, which includes the assimilation of all surface observations and employs the NSSL two-moment microphysics scheme, produced POD values ~ 20% higher compared to both experimental runs. Qualitative comparisons of reflectivity values and updraft speed between the real-time and Thompson runs are also presented. Lastly, the impacts of removing the METAR observations from the 16 May 2017 run are discussed.

1. INTRODUCTION

The National Weather Service (NWS) issues hundreds of hazardous weather warnings to the public every year. These advisories have improved spatially and temporally over the past 30 years, with lead times on tornado warnings increased from 3 minutes in 1978 to a current average of 13 minutes. Since 2006, lead times have stagnated at 13 minutes using the current warning paradigm (Stensrud et al. 2013). Thus, we are researching an enhanced system that will combine more accurate convective-scale forecasts with the current radar-based warning procedure.

The NOAA Warn-on-Forecast (WoF) research project aims to produce short-term, high resolution probabilistic forecasts via a regional, 1-km storm-scale ensemble prediction system. A prototype system, the NSSL Experimental Warn-on-Forecast System for ensembles (NEWS-e), employs a 36-member Weather Research and Forecasting (WRF) model ensemble with the assimilation of radar products, satellite retrievals, and surface observations. It has been shown that the assimilation of these datasets has led to

Corresponding author address:
Francesca Lappin, Florida State University
Center for Analysis and Prediction of Storms,
120 David L. Boren Blvd., Norman, OK 73072
Email: fml14@my.fsu.edu

improved prediction of severe storms and their environments (Wheatley et. al, 2015; Jones et al. 2016). The NEWS-e is run over an 8-hour period each day beginning at 1800 UTC with 3 hour forecasts produced at the top of the hour and 90 minute forecasts at the half hour, with 5-minute output time. This allows us to deliver very short range (0–3 hr) forecasts for various types of hazardous weather.

The configuration of the NEWS-e changed from spring 2016 to spring 2017 in an attempt to improve the system, which included a new microphysics scheme and the assimilation of surface observations from METAR. The NSSL microphysics is a two-moment scheme (Mansell et al. 2010) for cloud droplets, rain drops, ice crystals, snow, graupel, and hail, and is more computation expensive than the Thompson microphysics (Thompson et al. 2004, 2008), which only predicts the second moment for raindrops and ice crystals. Furthermore, previous studies have noted that the Thompson produces considerably greater quantities of snow compared to other schemes (Wheatley et al. 2014). Also, this year, assimilated surface observations were expanded to include METAR data. Prior to this year only Oklahoma Mesonet surface observations were assimilated into the NEWS-e. Consequently, if the grid were to fall outside of Oklahoma there would be no surface data assimilated into the model.

These changes warrant an investigation of how they impact the evolution of simulated storms in NEWS-e forecasts. The process of analyzing the changes and comparable performance of each run will be further explained in Section 2. The cases that the runs are modeled after will also be discussed in Section 2. Lastly, Section 3 will present comparisons of NEWS-e forecasts run with the NSSL two-moment scheme and Thompson microphysics scheme, as well as those with the METAR data removed.

2. METHODOLOGY

The primary goal of the WoF project is to increase tornado, high wind, and flash flood lead times. In this study, we examine the ability of NEWS-e forecasts to reproduce strongly rotating thunderstorms that occurred during several events. A brief summary of the methodology employed in this study follows below. This is attempted through the use of ensemble forecasts, allotting probabilities for event likelihood.

Three case days will be used to evaluate the differences associated with the changes to the NEWS-e configuration used in spring 2017. Two experiments will be conducted for each case. The first will remove the METAR surface observations from assimilation. The second will replace the NSSL two-moment scheme with the Thompson. Observational differences will then be noted from a side by side comparison and further evaluated statistically.

The spring 2017 NEWS-e configuration will be compared to the above experiments using various qualitative and quantitative approaches. For example, forecasted storm rotation objects (identified in 2-5 km updraft helicity fields) will be verified against observed storm rotation objects (identified in azimuthal wind shear fields) to be able to compare the relative performance (in a bulk sense) of forecasts from one experimental setup to another. Differences in storm tracks due to the addition of surface observations will be initially assessed through a member-by-member evaluation. Model fields such as reflectivity, vertical vorticity, and updraft will be compared using vertically averaged profiles. These vertical profiles will be calculated by averaging—at each model level—all grid points produced by all members at all output times in a given forecast. For each variable considered, a threshold is imposed to eliminate relatively small values. The threshold for reflectivity is 25 dBZ (most likely to indicate precipitation), .002 s^{-1} for vertical vorticity .002 s^{-1}, and 5 m $s^{-1}$for updraft speed. These thresholds were selected to isolate the areas of most significant convection and/or rotation, while mitigating the effects of spurious convection.

a. Case summaries

1) 9 May 2017

Around 1800 UTC convection initiated across central New Mexico that quickly grew into supercells which traveled across the state and into western Texas. The environment was characterized by low Convective Available Potential Energy (CAPE) of 500–1500 J/kg, with high bulk shear of 40–50 kts. A retreating dryline and moisture advection lends itself to the propagation of thunderstorms. Adding in the strengthening of a low-level jet leads to the development for mesocyclones and strengthening storm-relative helicity (SRH). On this day 11 tornadoes, 25 wind reports, and 29 hail reports

were produced in the area. This event was chosen due to the presence of supercell development in a low CAPE, high wind shear convective environment. In past studies, this type of environment has produced lower performance in NEWS-e forecasts (Jones et al. 2016).

2) 16 May 2017

Around 2000 UTC convection initiation occurred over western Oklahoma and the Texas panhandle. The convection intensified rapidly, evolving into supercells by 2100 UTC. One of the supercells became tornadic, producing a long-track EF3 tornado in Elk City, OK. Overall on this day, 34 tornados were reported with 99 wind and 115 hail reports as well. Pre-storm environmental conditions include large CAPE values, 3000–4000 J/kg, rapidly increasing SRH, and low-level moistening priming the area for convection. The sharpening of a dryline triggers the rapid production of supercells along the border. The presence of a distinctive dryline is part of the reason this case was chosen because of the rapid temperature and moisture change along the boundary. This warrants the possible influence of more surface data points. Similarly, the domain already includes the assimilation of Oklahoma mesonet data so the impact of including more surface data points to the model can be evaluated.

3) 23 May 2017

2100z marks the time of initiation for the collection of storms which lined up to cross central Texas and head to the southeast coast. This event was a primarily high-speed wind event with some hail reports. The environment was characterized by low CAPE, 500–1000 J/kg, and strengthening bulk shear values of 40–50 kts, conducive for hail and strong winds. The passage of a cold front triggered convective development along the leading edge. This case differs from the other two because it was primarily wind driven and lacked significant rotation. Also, since the grid does not fall within Oklahoma Mesonet range the inclusion of METAR data accounts for all the surface observations included in the model.

3. RESULTS

a) Quantitative Metrics

The false alarm ratio (FAR) and POD were calculated for each experiment and compared to the real time run. These metrics were calculated using object-based matching of rotational features from modeled 2–5 km UH with observed azimuthal wind shear. Azimuthal wind shear is used as a proxy for vertical vorticity since it can be calculated from radial velocity returns and is roughly half the vertical vorticity (Skinner et al. 2016).

Comparison of the POD scores for the 9 May 2017 real-time (RLT) versus the experimental METAR-removed (NOSFC) run shows that the POD values for the RLT run are ~0.2 higher until forecast minute 100 when the two scores become similar (Figure 1a). A similar evolution in the POD values is exhibited during the first 100 minutes of the forecast in the 16 May 2017 case with the POD scores higher by ~ 0.25 in the RLT run. However, unlike 9 May, the RLT run POD scores remain higher through the remainder of the forecast period (Figure 1a).

The POD values from 23 May 2017 have an odd evolution likely due to the type of event (damaging straight-line winds). POD scores in the RLT run are ~ 0.3 higher compared to the NOSFC run for the first 75 minutes of the forecast (Figure 3a). After this point, a wave-like pattern is shown, which is a stark difference from the other two cases. The reason for this evolution in POD values is still under investigation.

The POD score differences between the RLT run and the experimental run that uses the Thompson scheme (THOMP) are similar to the RLT versus NOSFC runs for the 9 May (Figures 1a, c) and 23 May (Figures 3a, c) events. The 16 May 2017 case exhibits the largest difference in POD scores between the NOSFC and THOMP experiments (Figures 2a, c).

The FAR values are generally high for the RLT, the NOSFC, and the THOMP runs. The RLT run from 23 May 2017 has FAR values 0.05 – 0.1 higher than both the NOSFC (Figure 3b) and THOMP (Figure 3d) runs. In addition, the FAR stays above 0.6 throughout all of the forecast times. The 9 May 2017 case (Figure 1b, d) exhibits a similar performance with FAR scores sitting above 0.6 throughout, but with little to no variability between either of the experimental runs or the RLT run. The 16 May 2017 case is the only one with FAR values below 0.5 (Figure 2b, d). Around forecast minute 65, the RLT run has a smaller FAR than the NOSFC run, ~ 0.1.

The RLT runs show considerable improvement in the POD when compared to either experiment over all case days, up to around 100

minutes into the forecast. However, there is no real improvement in the FAR for the RLT run. It is notable that the number of objects the RLT run produces is 57–84% greater than that of the experiments.

b) Rainfall Accumulation

Subjective comparisons of the ensemble-mean total precipitation fields produced by the THOMP and RLT runs show a high bias in the RLT run. Across the domain, the RLT run produces swaths of rain in the 0.5–1.5 in range, while the THOMP run produces areas in the 0.01–0.5 in range (Figure 4a, b). When compared to the observed Stage IV precipitation (not shown) both experimental runs smear together defined, observed precipitation tracks into one mass.

In a pool of roughly 41.6 million rainfall points, the difference between the number of points the THOMP and RLT produced is always less than 1% of the total. Nonetheless, there is a consistent pattern in how the difference in rainfall accumulation is distributed. In the <0.25 in bin the THOMP produces more points. This is consistent with the known tendency of the Thompson scheme to produce larger stratiform rain areas due to higher amounts of snow production within the cloud (Wheatley et al. 2014). For every other rainfall amount, the RLT run consistently produces more accounts of rain across all cases (Figure 5).

c) Vertical Profiles

Throughout the column, the reflectivity values produced in the THOMP and RLT runs illustrate a pattern seen across all cases and forecast hours. Below level 14 (4-5 km AGL), the RLT run has reflectivity values of 35–38 dBZ, with the THOMP run producing values 2–4 dBZ less than that. Between level 14 and 15, reflectivity values increase in the THOMP run by 2–3 dBZ, with the RLT run exhibiting a lesser rate of decline. Above level 14, the THOMP run has higher reflectivity values for the remainder of the vertical profile (Figure 6). This signature is present throughout all the cases studied at any forecast hour.

The change in 2–5 km updraft is comparable to that of the reflectivity profiles. Below level 14, the profiles from each run lie close to one another, indicating similar updraft speeds. Above level 14, the updraft velocities in the THOMP run continue to increase while those in the RLT run plateau. (Figure 7). Thus, the THOMP run has greater updraft velocities on average of ~ 2 m s^{-1} with the greatest difference of ~ 4 m s^{-1}. This pattern is representative of the majority of profiles.

d) 16 May 2017 RLT/NOSFC Comparison

The 16 May 2017 case was highlighted by two storm tracks over western Oklahoma, including the El Reno supercell. Some members from the NOSFC run show only one major track and fail to represent the El Reno supercell. This in part explains why the POD score is 0.25 lower for the NOSFC run yet the FAR is initially 0.15 lower than the RLT run. The portrayal of only one track occurs as well in the RLT run, but with less members. The exact cause of this is yet to be discovered, but it may be due to enhanced convergence and moisture pooling in the RLT run.

4. SUMMARY AND DISCUSSION

This study examines changes made to the NEWS-e configuration during spring 2017, and more specifically their impact(s) on NEWS-e forecasts of three severe weather events selected from spring 2017. The first experiment involved data-denial, whereby the METAR data was not assimilated during the retrospective run. The second experiment involved changing the microphysics scheme from the NSSL 2-moment (NEWS-e configuration for spring 2017) to the Thompson partial two-moment scheme.

In this study, for each event, forecasted storm rotation objects (identified in 2–5 km updraft helicity fields) were verified against observed storm rotation objects (identified in azimuthal wind shear fields), and it was found that the assimilation of surface observations from METAR produced forecasts with consistently higher POD scores, by as much as 20% at early forecast times. It was noted, at times, that the addition of METAR data allowed for both improvements in the location of convection initiation and subsequent supercell development. The change from a partial two-moment to a fully two-moment microphysics scheme produced similar improvements in the forecast POD scores.

While efforts are ongoing to identify the physical mechanisms responsible for the above scores, a couple of differences between the Thompson and NSSL 2-moment microphysics schemes have been already been shown through examination of vertically averaged profiles of

reflectivity and updraft speed. First, above ~4–5 km AGL we find that the Thompson scheme tends to produce higher values of reflectivity and increased updraft speeds. Below this level, the two schemes produce similar (average) values of updraft speed, while the NSSL 2-moment scheme produces relatively higher values of reflectivity.

A final distinction between the two microphysics schemes is revealed in accumulated rainfall plots. The Thompson scheme produces too great an area of rainfall amounts in the ~0.0–0.5-inch rainfall band. On the other hand, the NSSL 2-moment scheme tends to produce too much rainfall (relative to the Thompson scheme) for all values greater than ~0.5 inches.

Future work will place a greater emphasis on the data-denial runs, where the METAR surface data was not included in the assimilation procedure. Again, it was found in this study that the inclusion of more surface data generally produced higher POD scores. Moreover, we should perform this analysis across more cases to test the generality of these findings.

5. ACKNOWLEDGMENTS

The author would like to personally thank Dr. Daphne LaDue for the opportunity to be a part of the Real-World Research Experience for Undergraduates at the National Weather Center, as well as her support and devotion to the program. In addition, thanks to my colleagues in the program who have provided endless support.

This work was prepared by the authors with funding provided by National Science Foundation Grant No. AGS-1560419, and NOAA/Office of Oceanic and Atmospheric Research under NOAA-University of Oklahoma Cooperative Agreement #NA11OAR4320072, U.S. Department of Commerce. The statements, findings, conclusions, and recommendations are those of the author(s) and do not necessarily reflect the views of the National Science Foundation, NOAA, or the U.S. Department of Commerce.

6. REFERENCES

Jones, T.A., K. Knopfmeier, D. Wheatley, G. Creager, P. Minnis, and R. Palikonda, 2016: Storm-scale data assimilation and ensemble forecasting with the NSSL Experimental Warn-on-Forecast System. Part II: combine radar and satellite data experiments. *Wea. Forecasting,* **31**, 297-327. https://doi.org/10.1175/WAF-D-15-0107.1

Mansell, E.R., C.L. Ziegler, and E.C. Bruning, 2010: Simulated Electrification of a Small Thunderstorm with Two-Moment Bulk Microphysics. *J. Atmos. Sci.,* **67**, 171–194,https://doi.org/10.1175/2009JAS2965.1

Simmons, K.M., Sutter, D., 2011. Economic and societal impacts of tornadoes. Amer. Meteor. Soc. 282. pp.

Skinner, P.S., L.J. Wicker, D.M. Wheatley, and K.H. Knopfmeier, 2016: Application of two spatial verification methods to ensemble forecasts of Low-Level Rotation. *Wea. Forecasting,* **31**, 713–735, https://doi.org/10.1175/WAF-D-15-0129.1

Stensrud, D. J., and Coauthors, 2013: Progress and challenges with Warn-on-Forecast. *Atmos.Res.,***123**,2–16, doi:10.1016/j.atmosres.2012.04.004

Thompson, G., R.M. Rasmussen, and K. Manning, 2004: Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part I: Description and sensitivity analysis. *Mon. Wea. Rev.,* **132**, 519–542, https://doi.org/10.1175/1520-0493(2004)132<0519:EFOWPU>2.0.CO;2

Thompson, G., P.R. Field, R.M. Rasmussen, and W.D. Hall, 2008: Explicit Forecasts of Winter Precipitation Using an Improved Bulk Microphysics Scheme. Part II: Implementation of a New Snow Parameterization. *Mon. Wea. Rev.,* **136**, 5095–5115, https://doi.org/10.1175/2008MWR2387.1

Wheatley, D.M., N. Yussouf, and D.J. Stensrud, 2014: Ensemble Kalman filter analyses and forecasts of a severe mesoscale convective system using different choices of microphysics schemes. Mon. Wea. Rev., 142, 3243–3263, https://doi.org/10.1175/MWR-D-13-00260.1

Wheatley, D.M., K.H. Knopfmeier, T.A. Jones, and G.J. Creager, 2015: Storm-scale data assimilation and ensemble forecasting with the NSSL Experimental Warn-on-Forecast System. Part I: radar data experiments. *Wea. Forecasting,* **30**, 1795–1817, https://doi.org/10.1175/WAF-D-15-0043.1

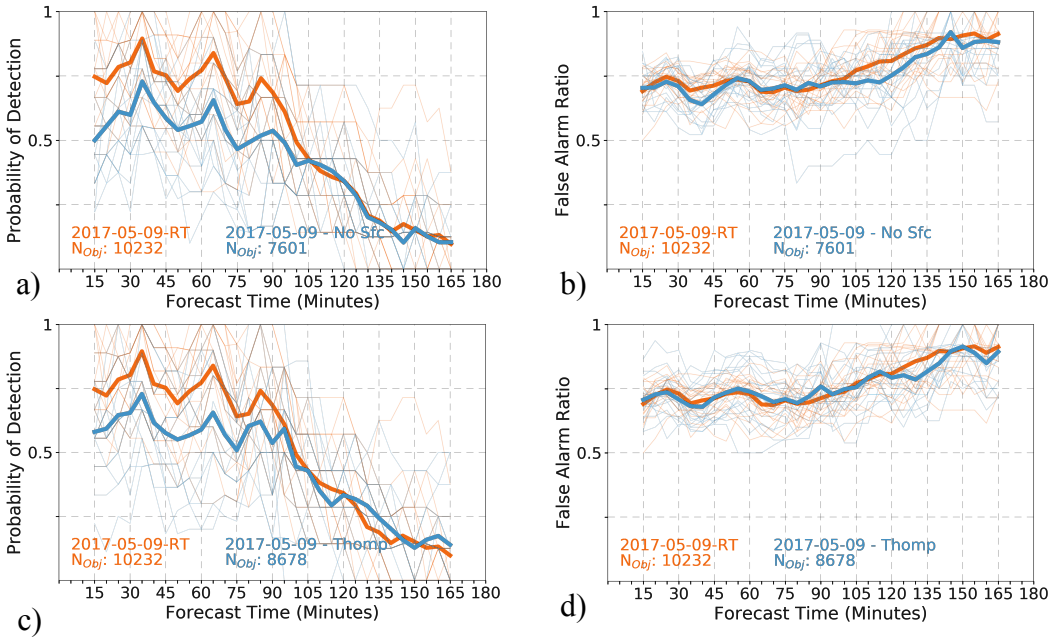

Fig. 1. (a) Probability of Detection (POD) for the NOSFC and RLT runs, (b) False Alarm Rate (FAR) for the NOSFC and RLT runs, (c) POD for the THOMP and RLT runs, and (d) FAR for the THOMP and RLT runs from all forecasts on 9 May 2017.

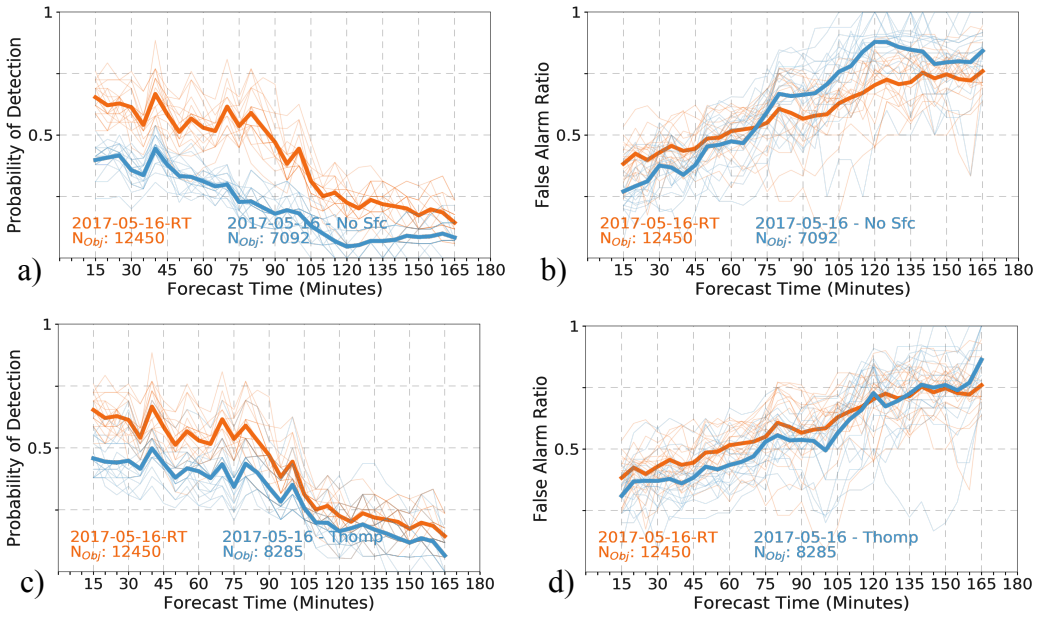

Fig. 2. Same as in Fig. 1, but for 16 May 2017.

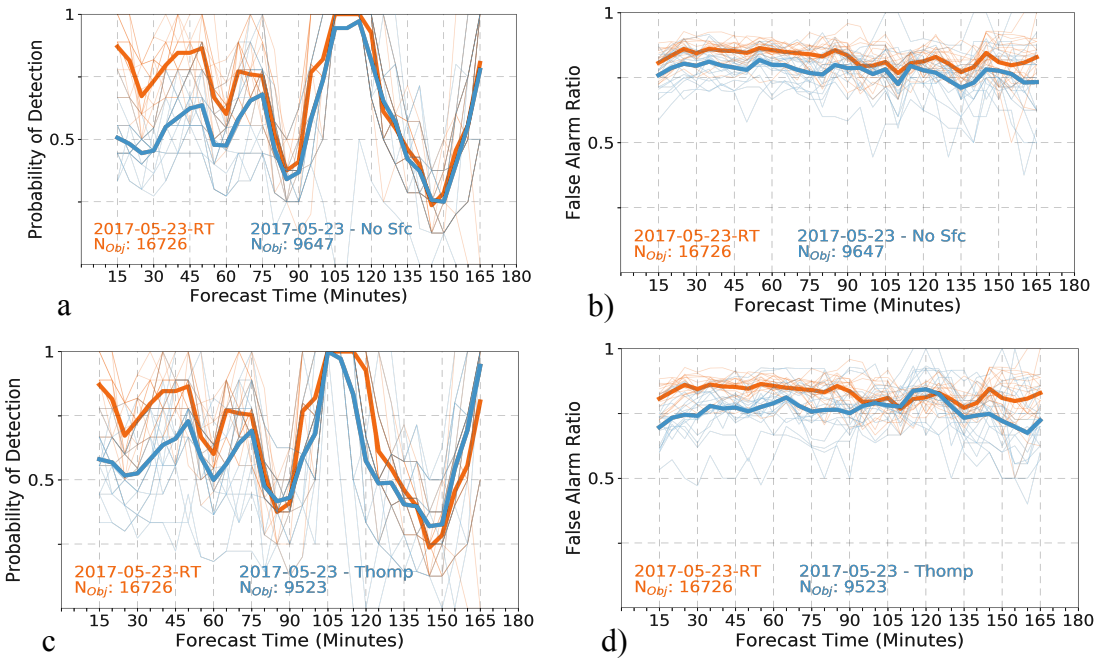

Fig. 3. Same as in Fig. 1, but for 23 May 2017.

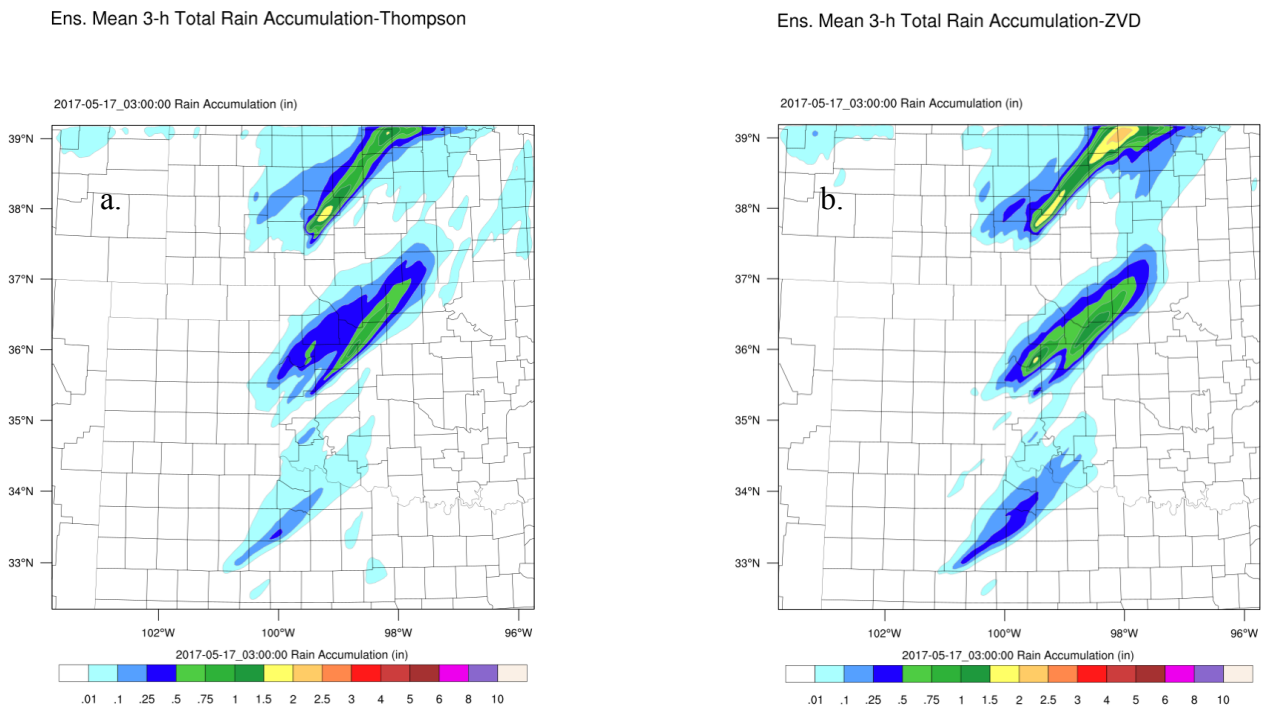

Fig. 4. 16 May 2017 total 3-hr rain accumulation (in) field modeled using the (a) Thompson scheme; (b) NSSL two-moment (RLT run)

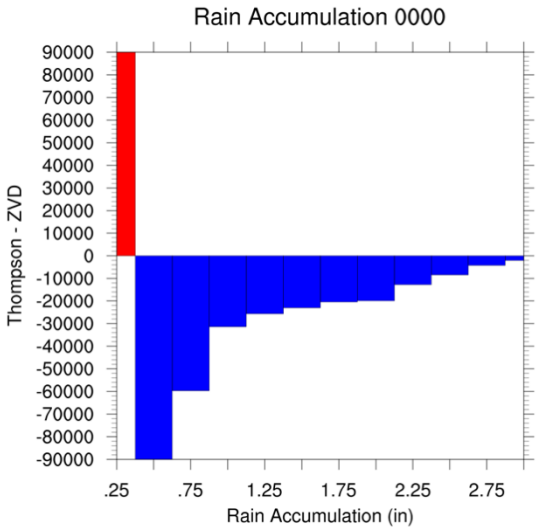

Fig. 5. Histogram of the difference between the experimental Thompson and real-time NEWS-e run rainfall accumulation (in) over a 3-hr forecast period ending at 0000 UTC 17 May 2017. Bins are every 0.25 in from 0.25 – 3.00 in.

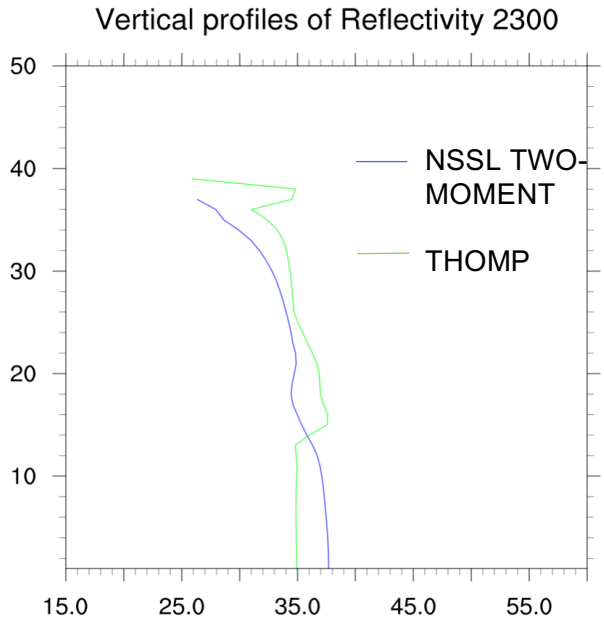

Fig. 6. Vertical profile of reflectivity (dBZ) averaged at each model level across the experimental domain from the 2300 UTC 16 May 2017 forecast. The real-time (RLT) run is depicted by the blue solid line, while the experimental Thompson (THOMP) run is indicated by the green line.

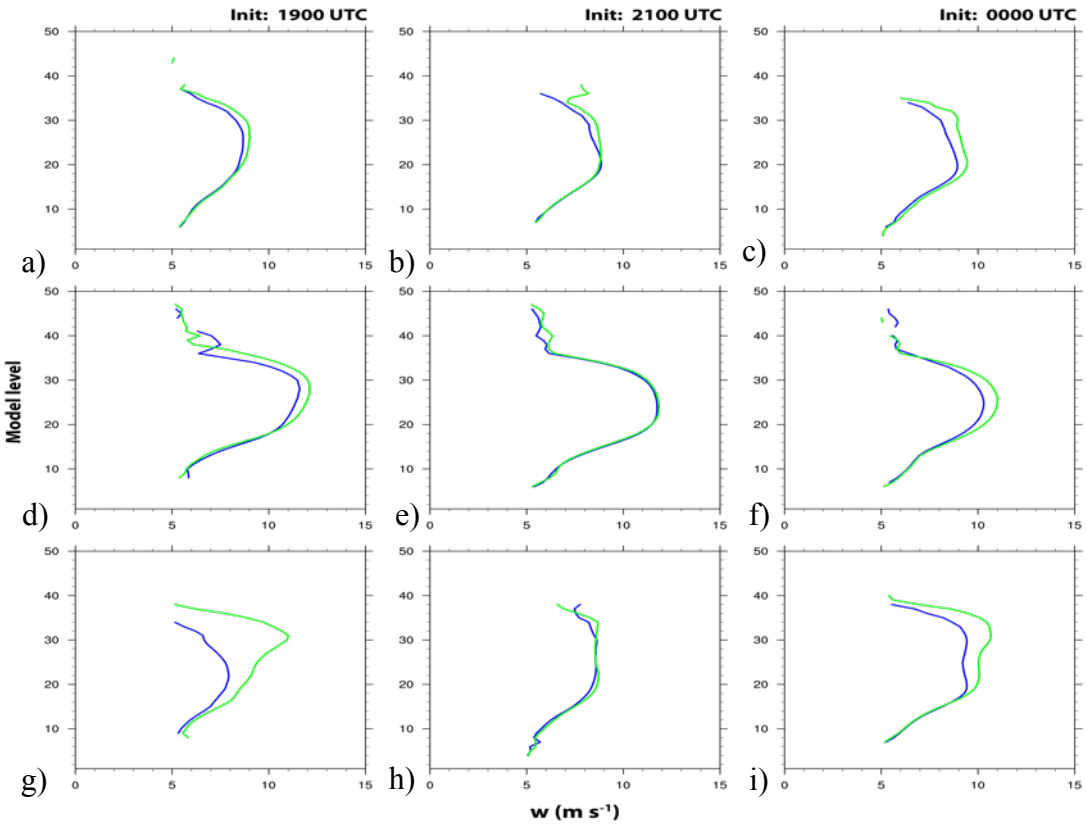

Fig. 7. Vertical profiles of vertical velocity averaged at each model level over the experimental domain for the 9 May 2017 case at a) 1900 UTC, b) 2100 UTC, and c) 0000 UTC 10 May 2017. d) – f) Same as in a) – c), but for the 16 May 2017 case. g) – i) Same as in a) – c), but for the 23 May 2017 case. Blue line is from the real-time NEWS-e run. Green line is from the experimental run using the Thompson microphysics scheme.